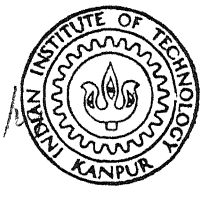
ON LONG PRIMITIVE, NARROW-SENSE BCH CODES

By

RAJINDER ARORA



DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

JANUARY, 1982

ON LONG PRIMITIVE, NARROW-SENSE BCH CODES

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of

MASTER OF PHILOSOPHY

By

RAJINDER ARORA

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

JANUARY, 1982

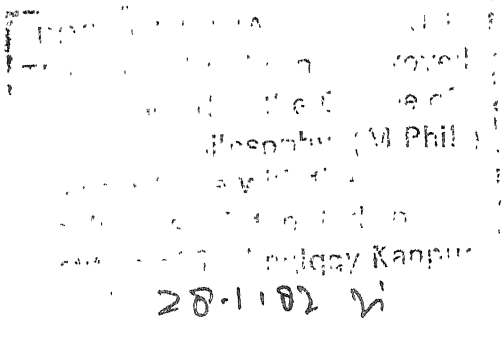
CERTIFICATE

This is to certify that the work embodied in this thesis entitled '' On Long Primitive, Narrow-Sense BCH Codes'' by Rajinder Arora has been carried out under my supervision and that it has not been submitted elsewhere for the award of any degree or diploma.

M C Bhandari

[M. C. BHANDARI]
Department of Mathematics
Indian Institute of Technology, Kan

January 7 , 1982.



I wish to express my sincere gratitude and regards to Dr. M. C. Bhandari for his guidance and valuable suggestions throughout the course of this work. I am deeply grateful, to all my friends, who have made this period of study a pleasant and memorable one.

Final thanks go to Mr. S. S. Pethkar and Mr. A. K. Bhatia for their excellent typing work.

राजेन्द्र

JANUARY 7 , 1982. [RAJINDER ARORA]

ABSTRACT

In this thesis we have proved that long primitive, narrow-sense BCH codes over GF(q) have distance $d \sim \frac{qn \ln R^{-1}}{\log_q n}$ thereby proving that long BCH codes are bad.

These results are based on the ''Enumeration of Information symbols in BCH Codes'', which gives an exact expression for rates of any sequence of primitive BCH codes of increasing length and fixed ratio of distance/length.

CONTENTS

1. INTRODUCTION

Coding theory deals with the problems of communication-that of encoding and decoding digital information for reliable transmission through communication channels with noise. One wishes to transmit a message which consists of a finite sequence of alphanumeric characters but due to imperfections in the communication channel a given transmitted character may be incorrectly received by the receiver. One can reduce the probability of error that a given transmitted character will be incorrectly received by the receiver by various encoding and decoding schemes. The basic idea behind these encoding and decoding schemes is as follows.

Suppose that we wish to transmit a sequence of binary digits across a noisy channel. Due to the presence of noise, a transmitted zero may be received as one and viceversa. Although we are unable to prevent the channel from causing such errors, we can however reduce their undesirable effects with the use of various encoding and decoding schemes. e.g., if a message of k-digits is to be transmitted, one adds certain number, say r of check digits to the message digits and transmits the entire block of n(=k+r) digits. Assuming that the channel noise changes few of these n transmitted digits, the r check digits may provide the receiver with the sufficient information to enable him to detect and/or to correct the channel errors. The rule which enables us to determine r check digits is called an encoding scheme.

A sequence of n-digits which the

encoder transmits is called a codeword. As r check digits are determined by the k-message digits, there will be 2^k codewords. The set consisting of these 2^k codewords is called the code. The rule which enable us to determine which of 2^k codewords was transmitted when any of the 2^n binary n-tuple is received is called a coding scheme.

Coding theory, in particular, has its origin in Shannon's (1948) famous theorem which guarantees the existence of codes that can transmit information at rates close to capacity with vanishingly small probability of error. In the past 33 years since Shannon's paper, coding theory has progressed rather fitfully through periods of empiric highs with discoveries of promising code classes, elegant decoding algorithms and dismal lows when it was feared that coding approach would never move beyond Golay [6] and Hamming [ ] codes. It now appears safe to say that coding is maturing into an important segment of communication engineering, one that will see increasing applic of relatively standard techniques to reduce system cost and complexity.

In the three years immediately following Shannon's paper, the highly significant work of Golay [6] and Hamming [ ] was published. The constructions of their codes have had a lasting influence on coding theory. In most of this early work the concept of pairty check equations was predominant. The idea of viewing a code as a subgroup of the Abelian group of 2^n bina n-tuples appears in the work of Slepian [14]. Moreover, Gol [6] certainly recognized the significance of working over the

integers modulo a prime. By 1961, the class of linear codes (also called group codes) was well established. The work of Slepian [14] provided a solid theoritical base for the investigation of such codes. The cyclic codes was introduced by Prange in 1957. In the year (1959), Hocquenghem, Bose and Chaudhuri [4] independently obtained multiple error correcting codes for the binary case. These codes are known as Bose-Chaudhu Hocquenghem (BCH) codes. The cyclic nature of BCH codes was firs shown by Peterson [11].

The discovery of efficient decoding algorithms for BCH codes by Peterson [11], Gorenstein and Zierler [7] and Berlekamp (1966) has emphasized the importance of these codes. The algorithms associate each coordinate place in a codeword with an element of a Galois field, which reduces the decoding operati to solving a set of equations over a finite field. Such an operation requires considerably less computation than the word by word search for the codeword closest to the received word.

An important result that long BCH codes are bad was first established by Lin and Weldon Jr. [10]. In view of the above result, in order to find very long, powerful, random error correcting codes, one must look elsewhere. But BCH codes of moderate length are rather powerful random error correcting codes.

First result on the number of information symbols in BCH codes came from Henry B. Mann [9]. However, he proved his results under a condition on the designed distance of the codes. This was studied in the general form by E.R. Berlekamp [1]. In 1972, Berlekamp [2] has proved that long primitive, binary BCH codes have distance $d \sim 2n\ln R^{-1}/\log_2 n$ where R is the information rate and n is the block length. He also obtained the upper and lower bounds on the designed and actual distances of any sequence of extended primitive BCH codes of increasing length n and fixed rate R. This has been generalized to BCH codes over any finite field in chapter 4. Our main result is "Long Primitive, Narrow-sense BCH Codes Over GF(q) have distance $d \sim q n \ln R^{-1}/\log_q n$".

Chapter 2, presents a brief account of preliminary results on finite fields, linear, cyclic and BCH codes, that are needed for the work done in the succeeding chapters. We have also discussed in brief the decoding procedure for linear codes. An example of BCH codes is worked out in detail to make the concept more clear. Chapter 3, present important results on the enumeration of information symbols in BCH codes. In the end of chapter 4, we have given the achievements of this work. Certain related directions for further studies are suggested.

2. [illegible]

Fields : A non empty set F equipped with two binary operations usually called addition and multiplication (denoted by + and . respectively) is called a field if (1) (F, +) is an abelian group, (2) $(F \setminus \{o\}, .)$ is an abelian group and (3) a.(b+c) = a.b + a.c for all a,b,c in F. If K is a field containing F as a subfield then K is said to be an extension of F. The prime subfield of K is the intersection of all subfields of K and hence is the smallest subfield of K. Prime subfield of a given field is either isomorphic to the field of rational numbers or integers modulo p where p is a prime. In the first case the characteristic of the field will be zero while p in the later case.

Let F be a field. Then F[X], the ring of all polynomials over F in the variable X is a principal ideal domain. If f(X) is an irreducible polynomial in F[X] then there exist a smallest extension K, isomorphic to the quotient field F[X]/(f(X)), of F which contains a zero of f(X). For any polynomial g(X), (g(X)) denotes the ideal in F[X] generated by g(X). If the irreducible polynomial f(X) is of degree n, then the degree of K over F(denoted by [K:F]) is n. Infact if a is the zero of f(X) in question then K = F(a) and the elements $1, a, \ldots, a^{n-1}$ form a basis for K over F as a vector space.

A Galois field is a finite field, i.e., a field having finite number of elements. Any finite field with characteristic p has p^n elements for some positive integer n and upto isomorphism there is only one field with p^n elements. A finite field having q elements is often denoted by GF(q). Nonzero elements of a

finite field form a cyclic group under multiplication. An element of the finite field is called a primitive element if it is the generator of the cyclic group.

If $f(X)$ is an irreducible polynomial of degree n with coefficients in $GF(p)$ ($\cong Z_p$) then $GF(p)[X]/(f(X))$ is a field having p^n elements. In fact

$$GF(p)[X]/(f(X)) = \left\{ r(X) + (f(X)) \,:\, r(X) \in GF(p)[X],\ \deg r(X) < \deg f(X) \text{ or } r(X) = 0 \right\}$$

$$= \left\{ r(X) + (f(X)) \,:\, r(X) = c_1 X^{n-1} + c_2 X^{n-2} + \ldots + c_n,\ c_i \in GF(p),\ i = 1, 2, \ldots n \right\}$$

Hence upto isomorphism $GF(p)[X]/(f(X))$ can be thought of as the set $\left\{ (c_1, c_2, \ldots, c_n) \,:\, c_i \in GF(p),\ i = 1, 2, \ldots, n \right\}$ of all n-tuples with the algebraic operations induced by the isomorphism.

Example 1 : The Galois field of $27 = 3^3$ elements. Let $GF(3) = \{0, 1, 2\}$. Then $f(X) = X^3 + X^2 + 2X + 1$ is an irreducible polynomial of degree 3 over $GF(3)$ and hence $GF(3)[X]/(f(X))$ is a field having $27 = 3^3$ elements. If γ is a root of $f(x)$ then $\gamma^3 + \gamma^2 + 2\gamma + 1 = 0$ or $\gamma^3 = 2\gamma^2 + \gamma + 2$. Thus the elements of $GF(27)$ are $0, 1, \gamma, \gamma^2, \ldots, \gamma^{25}$ where

2. [illegible]

Fields : A non empty set F equipped with two binary operations usually called addition and multiplication (denoted by + and . respectively) is called a field if (1) (F, +) is an abelian group, (2) (F $\setminus$ {o} , .) is an abelian group and (3) a.(b+c) = a.b + a.c for all a,b,c in F. If K is a field containing F as a subfield then K is said to be an extension of F. The prime subfield of K is the intersection of all subfields of K and hence is the smallest subfield of K. Prime subfield of a given field is either isomorphic to the field of rational numbers or integers modulo p where p is a prime. In the first case the characteristic of the field will be zero while p in the later case.

Let F be a field. Then F[X], the ring of all polynomials over F in the variable X is a principal ideal domain. If f(X) is an irreducible polynomial in F[X] then there exist a smallest extension K, isomorphic to the quotient field F[X]/(f(X)), of F which contains a zero of f(X). For any polynomial g(X), (g(X)) denotes the ideal in F[X] generated by g(X). If the irreducible polynomial f(X) is of degree n, then the degree of K over F(denoted by [K:F]) is n. Infact if a is the zero of f(X) in question then K = F(a) and the elements $1, a, \ldots, a^{n-1}$ form a basis for K over F as a vector space.

A Galois field is a finite field, i.e., a field having finite number of elements. Any finite field with characteristic p has p^n elements for some positive integer n and upto isomorphism there is only one field with p^n elements. A finite field having q elements is often denoted by GF(q). Nonzero elements of a

finite field form a cyclic group under multiplication. An element of the finite field is called a primitive element if it is the generator of the cyclic group.

If $f(X)$ is an irreducible polynomial of degree n with coefficients in $GF(p)$ ($\cong Z_p$) then $GF(p)[X]/(f(X))$ is a field having p^n elements. In fact

$$GF(p)[X]/(f(X)) = \left\{ r(X) + (f(X)) \,:\, r(X) \in GF(p)[X],\ \deg r(X) < \deg f(X) \text{ or } r(X) = 0 \right\}$$

$$= \left\{ r(X) + (f(X)) \,:\, r(X) = c_1 X^{n-1} + c_2 X^{n-2} + \ldots + c_n,\ c_i \in GF(p),\ i = 1,2,\ldots n \right\}$$

Hence upto isomorphism $GF(p)[X]/(f(X))$ can be thought of as the set $\left\{ (c_1, c_2, \ldots, c_n) \,:\, c_i \in GF(p),\ i = 1,2,\ldots,n \right\}$ of all n-tuples with the algebraic operations induced by the isomorphism.

Example 1 : The Galois field of $27 = 3^3$ elements. Let $GF(3) = \left\{0,1,2\right\}$. Then $f(X) = X^3+X^2+2X+1$ is an irreducible polynomial of degree 3 over $GF(3)$ and hence $GF(3)[X]/(f(X))$ is a field having $27 = 3^3$ elements. If γ is a root of $f(x)$ then $\gamma^3 + \gamma^2 + 2\gamma + 1 = 0$ or $\gamma^3 = 2\gamma^2 + \gamma + 2$. Thus the elements of $GF(27)$ are $0, 1, \gamma, \gamma^2, \ldots, \gamma^{25}$ where

$0 = 0\gamma^2+0\gamma+0 \leftrightarrow (0,0,0)$, $1 = 0\gamma^2+0\gamma+1 \leftrightarrow (0,0,1)$

$\gamma = 0\gamma^2+\gamma+0 \leftrightarrow (0,1,0)$, $\gamma^2 = \gamma^2+0\gamma+0 \leftrightarrow (1,0,0)$

$\gamma^3 = 2\gamma^2+\gamma+2 \leftrightarrow (2,1,2)$, $\gamma^4 = 2\gamma^2+\gamma+1 \leftrightarrow (2,1,1)$

$\gamma^5 = 2\gamma^2+0\gamma+1 \leftrightarrow (2,0,1)$, $\gamma^6 = \gamma^2+0\gamma+1 \leftrightarrow (1,0,1)$

$\gamma^7 = 2\gamma^2+2\gamma+2 \leftrightarrow (2,2,2)$, $\gamma = 0\gamma^2+\gamma+1 \leftrightarrow (0,1,1)$

$\gamma^9 = \gamma^2+\gamma+0 \leftrightarrow (1,1,0)$, $\gamma^{10} = 0\gamma^2+\gamma+2 \leftrightarrow (0,1,2)$

$\gamma^{11} = \gamma^2+2\gamma+0 \leftrightarrow (1,2,0)$, $\gamma^{12} = \gamma^2+\gamma+2 \leftrightarrow (1,1,2)$

$\gamma^{13} = 0\gamma^2+0\gamma+2 \leftrightarrow (0,0,2)$, $\gamma^{14} = 0\gamma^2+2\gamma+0 \leftrightarrow (0,2,0)$

$\gamma^{15} = 2\gamma^2+0\gamma+0 \leftrightarrow (2,0,0)$, $\gamma^{16} = \gamma^2+2\gamma+1 \leftrightarrow (1,2,1)$

$\gamma^{17} = \gamma^2+2\gamma+2 \leftrightarrow (1,2,2)$, $\gamma^{18} = \gamma^2+0\gamma+2 \leftrightarrow (1,0,2)$

$\gamma^{19} = 2\gamma^2+0\gamma+2 \leftrightarrow (2,0,2)$, $\gamma^{20} = \gamma^2+\gamma+1 \leftrightarrow (1,1,1)$

$\gamma^{21} = 0\gamma^2+2\gamma+2 \leftrightarrow (0,2,2)$, $\gamma^{22} = 2\gamma^2+2\gamma+0 \leftrightarrow (2,2,0)$

$\gamma^{23} = 0\gamma^2+2\gamma+1 \leftrightarrow (0,2,1)$, $\gamma^{24} = 2\gamma^2+\gamma+0 \leftrightarrow (2,1,0)$

$\gamma^{25} = 2\gamma^2+2\gamma+1 \leftrightarrow (2,2,1)$.

Linear Codes : There are fundamentally two different type of codes block codes and tree codes. The encoder for a block code breaks the continuous sequence of information digits into k-symbol section or blocks. It then operate on these blocks independently. While a tree code operates on the information sequence without breaking it up into independent blocks. Throughout the thesis, we will consider only block codes. The work of Slepian [14] was the first investigation of general

ure into the problem of block coding. He studied a special
ass of codes called linear block codes.

finition 1 : A linear block code over GF(q) of block length n
d dimension k is a k-dimensional subspace C of $V_n(q)$ where $V_n(q)$
presents the space of all n-tuples over the field GF(q). C is
lled an (n,k) code. The elements of C are called codewords.

Let $u,v \in V_n(q)$. The Hamming distance $d(u,v)$ between u and v
s the number of coordinate places in which they differ and the
Hamming weight $w(u)$ of u is the number of nonzero coordinates of u.
he minimum distance of a linear block code C is defined by

$$d = \min_{\substack{u \neq v \\ u,v \in C}} \{ d(u,v) \}$$

Let C be a linear block code. If $u,v \in C$ then $d(u,v) = d(u-v,o)$
$w(u-v)$. Hence the minimum distance of C is the weight of a nonzero
odeword having least weight. Let $v_1, v_2, \ldots, v_k$ be a basis for C.
hen any linear combination $\sum_{i=1}^{k} \gamma_i v_i$, $\gamma_i \in GF(q)$ is a codeword.
n otherwords if G is the matrix whose rows are $v_1, v_2, \ldots, v_k$ then
is the row space of G. G is called the generator matrix of the code.

An orthognol complement $C^{\perp}$ of C is an (n-k)-dimensional subspace
f $V_n(q)$ and hence a (n,n-k) code $C^{\perp}$ is called the dual code of C.
et H be the generator matrix of $C^{\perp}$ then

$$GH^T = o$$

where H^T stands for the transpose of the matrix H. H is called the parity check matrix of the code C.

Let H be any $n \times n$ matrix over $GF(q)$ of rank $n-k$ and $C = \{v : vH^T = o\}$ then C is k-dimensional subspace of $V_n(q)$ and hence is an (n,k) code.

Decoding of a linear block code C : Consider the factor space $V_n(q)/C$. Write each coset as the row of an array. The first row of the array contains the elements of C and we choose the first element to be the zero vector. We place in each row at the first place, the elements having the least weight. We denote this array by $S = (S_{ij})$ where each S_{ij} is an element of $V_n(q)$. Note that S is an $q^{n-k} \times q^k$ matrix over $V_n(q)$. Let the code vectors be $C_1 = o, C_2, \ldots, C_{q^k}$ and $x_1, x_2, \ldots, x_{q^{n-k}}$ are the leaders of the cosets, i.e., minimal weight element of the cosets. Then $S_{ij} = x_i + C_j$. The above array S is called Slepian's standard array. A decoding rule is that if the received codeword is $S_{ij} = x_i + C_j$ then the transmitted codeword is C_j. It is easy to see that if all possible error patterns with t or fewer errors occur as coset leaders, then this decoding will correct all patterns of t or fewer errors. The following two theorems summerize the results from [7].

Theorem 1[7] : A linear code C is t-error correcting if and only if wt(X) > 2t for every nonzero X in C.

Theorem 2[7] : If H is the pairty check matrix of a linear code C of length n, then the code has minimum distance d if an only if every d-1 columns of H are linearly independent and some d columns of H are linearly dependent.

Cyclic Codes : The class of linear codes is still too general for us to say much on the problem of constructing codes with good distance properties. An important class of codes known as cyclic codes was studied by Prange (1957). Suppose $v = (a_0, a_1, \ldots, a_{n-1})$ is an element of $V_n(q)$ then a cyclic shift of the coordinates of v is the element $(a_{n-1}, a_0, \ldots, a_{n-2})$ in $V_n(q)$.

Definition 2 : A linear code is said to be cyclic if every cyclic shift of the coordinates of a codeword is again a codeword.

As we know GF(q)[X], the set of all polynomials over GF(q) in the variable X is a principle ideal domain. X^{n-1} is a polynomial in GF(q)[X]. Then GF(q)[X]/(X^n-1) is an n-dimensional algebra, denoted by A_n over GF(q). There is a natural isomorphism between A_n and the algebra of all polynomials over GF(q) of degree less than n with multiplication defined modulo (X^n-1). Therefore in view of the above isomorphism each element of A_n is represented by a polynomial of degree less than n. Using this representation

the following theorem gives a necessary and sufficient condition for a linear code to be a cyclic code.

Theorem 3 : An (n,k) linear code C over $GF(q)$ is cyclic if and only if it is an ideal of A_n.

Proof. Let C be an ideal of A_n. It is clearly a linear subspace of A_n and it remains to check that it is a cyclic space. Let $v \in C$. Then $v = a_0X^{n-1} + a_1X^{n-2}+\ldots+a_{n-1}$. Since C is an ideal of A_n, $X.(a_0X^{n-1}+a_1X^{n-2}+\ldots+a_{n-1}) = a_0X^n+a_1X^{n-1}+\ldots+a_{n-1}X = a_1X^{n-1}+\ldots+a_{n-1}X+a_0$ (modulo X^{n-1}) is an element of C. But $a_1X^{n-1}+\ldots+a_{n-1}X+a_0$ is a cyclic shift of v.

Conversely let C be a cyclic subspace. Let $v \in C$ and let $a_0X^{n-1}+\ldots+a_{n-1}$ be any element of A_n. As $v \in C$ and C is a cyclic subspace , Xv and hence X^jv is an element of C for positive integer j. Hence for any element $a_0X^{n-1}+a_1X^{n-2}+\ldots+a_{n-1} \in A_n$, and $v \in C$, $(a_0X^{n-1}+a_1X^{n-2}+\ldots+a_{n-1}).\ v = a_0X^{n-1}v + a_1X^{n-2}v + \ldots + a_{n-1}v \in C$. Therefore C is an ideal of A_n.

The polynomial X^n-1 plays a key role in the theory of cyclic codes. Consider $f(X) = X^n-1$, $f'(X) = nX^{n-1}$. Let 'a' be a root of $f(X)$. Then $a^n-1 = 0$ and so a is nonzero. Hence $f'(a) \neq 0$ if and only if n and q are relatively prime (i.e., $(n,q) = 1$). Thus a necessary and sufficient condition that X^n-1 does not have any factor

of multiplicity greater than one over GF(q) is that (n,q) = 1. The following theorem characterizes the ideals of A_n.

Theorem 4 : The unique monic polynomial g(X) of minimal degree in any ideal A of A_n is a generator of A and divides X^n-1. The dimension of A is the degree of g(X). Conversely, a divisor of X^n-1 is a generator of an ideal in A_n.

Proof : Let g(X) be the unique monic polynomial of minimal degree in A and let f(X) be any element in the ideal. By division algorithm f(X) = q(X)g(X)+r(X) where degree of r(X) < degree of g(X). Since A is an ideal, r(X) = f(X)-q(X)g(X) is in A. Since g(X) is of minimal degree in A, r(X) = o. Similarly taking $f(X) = X^n-1$, it follows that g(X) divides X^n-1. Suppose g(X) is of degree n-k. Then the elements g(X), Xg(X), $X^2g(X)$,.... $X^{k-1}g(X)$ are linearly independent in A. Since any element in the ideal is of the form a(X)g(X) where a(X) is of degree less than k, these elements span the ideal . Thus A has dimension k.

Conversely suppose that g(X) is any monic polynomial that divides X^n-1 and consider $A_n g(X)$. Let $f(X) \in A_n g(X)$. Then f(X) = a(X)g(X) mod(X^n-1) = a(X)g(X)+b(X)(X^n-1)= p(X)g(X) as g(X) divides X^n-1. Hence g(X) is the required generator of $A_n g(X)$.

The polynomial $h(X) = X^n-1 \;/\; g(X)$ is called the check polynomial of the cyclic code. If $g(X) = \sum_{i=0}^{n-k} g_i X^i$ than the generator matrix of the code has the form

$$G = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k-1} & 1 & 0 \ldots 0 & 0 \\ 0 & g_0 & g_1 & \cdots & g_{n-k-2} & g_{n-k-1} & 1 \ldots 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \cdots \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \cdots \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \cdots \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 0 & g_0 & g_1 \cdots g_{n-k-1} & 1 \end{bmatrix}$$

An alternative specification of cyclic codes can be made in terms of the roots, possibly in an extension field, of the generator $g(X)$ of the ideal. Let $\gamma_1, \gamma_2, \ldots, \gamma_r$ be all the roots of $g(X)$. We assume that there are no repeated roots. Then the following statement uniquely specifies a cyclic code : A vector $f(X)$ is codeword if and only if $\gamma_1, \gamma_2, \ldots, \gamma_r$ are the roots of $f(X)$ since in that case $g(X)$ divides $f(X)$. If the minimum function of γ_i is $m_i(X)$ then $f(X)$ is a cod vector if and only if $f(X)$ is divisible by $m_1(X), m_2(X), \ldots, m_r(X)$ and hence by their least common multiple. Now $g(X)$ is the monic polynomial of minimum degree having $\gamma_1, \gamma_2, \ldots, \gamma_r$ as the roots. Therefore, the code is the ideal Generated by

$g(x) = LCM\ [\ m_1(x), m_2(x), \ldots, m_r(x)\ ]$.

Hamming Codes : The original codes of Hamming [8] were restricted to the binary case. Define $m \times (2^m - 1)$ pairty check matrix H for these codes by taking all nonzero binary m-tuples as the columns. The addition of any two columns of H cannot be zero. Hence every two columns of H are linearly independent . However it is easy to find three linearly dependent columns since the addition of two distinct binary m-tuples is again a nonzero binary m-tuple . Thus by theorem 2, the pairty check matrix H defines a code with distance three and hence is one-error correcting code. The rank of H is clearly m and so this is an $(2^m - 1, 2^m - m - 1)$ code.

In the generalized Hamming codes over GF(q), one defines the pairty check matrix H by taking all nonzero m-tuples over GF(q) with the property that no two of its columns are scalar multiples of one another. It will be an mxn matrix when $n = q^m - 1/q - 1$. This matrix has rank m. Clearly every two columns of H are linearly independent and it is easy to find three linearly dependent columns. Thus by theorem 2, H defines a code with distance three. The result is a one-error-correcting $(n, n-m)$ code over GF(q) wheren $n = q^m - 1/q - 1$. The following theorem gives a lower bound on the minimum distance of cyclic codes.

Theorem 5 : Let g(x) be the generator polynomial of a cyclic code of length n over GF(q) and let $\gamma^{e_1}, \ldots, \gamma^{e_{n-k}}$ be the roots of g(x), possibly in an extension field where γ is an element of order n.

The minimum distance of the code is the largest number of consecutive integers modulo n in the set $S = \{e_1, e_2, \ldots, e_{n-k}\}$.

Proof : Let $m, m+1, \ldots, m+d-2$ denote the largest set of consecutive integers modulo n in the set S. A cyclic code with the roots $\gamma^{e_1}, \gamma^{e_2}, \ldots, \gamma^{e_{n-k}}$ is the null space of the matrix

$$H = \begin{bmatrix} (\gamma^{e_1})^{n-1} & (\gamma^{e_1})^{n-2} & \cdots & \gamma^{e_1} & 1 \\ (\gamma^{e_2})^{n-1} & (\gamma^{e_2})^{n-2} & \cdots & \gamma^{e_2} & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ (\gamma^{e_{n-k}})^{n-1} & (\gamma^{e_{n-k}})^{n-2} & \cdots & \gamma^{e_{n-k}} & 1 \end{bmatrix}$$

Consider the matrix

$$H' = \begin{bmatrix} (\gamma^{m})^{n-1} & (\gamma^{m})^{n-2} & \cdots & \gamma^{m} & 1 \\ (\gamma^{m+1})^{n-1} & (\gamma^{m+1})^{n-2} & \cdots & \gamma^{m+1} & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ (\gamma^{m+d-2})^{n-1} & (\gamma^{m+d-2})^{n-2} & \cdots & \gamma^{m+d-2} & 1 \end{bmatrix}$$

Then no linear combination of d-1 columns of the matrix H' is zero as for any set of d-1 of its columns,

$$\Delta = \begin{vmatrix} (\gamma^m, j_{d-1} & (\gamma^m, j_{d-2} & \cdots & (\gamma^m, j_1 \\ (\gamma^{m+1}, j_{d-1} & (\gamma^{m+1}, j_{d-2} & \cdots & (\gamma^{m+1}, j_1 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ (\gamma^{m+d-2}, j_{d-1} & (\gamma^{m+d-2}, j_{d-2} & \cdots & (\gamma^{m+d-2}, j_1 \end{vmatrix}$$

$$= \gamma^{m(j_1+j_2+\cdots+j_{d-1})} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \gamma^{j_{d-1}} & \gamma^{j_{d-2}} & \cdots & \gamma^{j_1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ (\gamma^{j_{d-1}})^{d-2} & (\gamma^{j_{d-2}})^{d-2} & \cdots & (\gamma^{j_1})^{d-2} \end{vmatrix}$$

$$= \gamma^{m(j_1+j_2+\cdots+j_{d-1})} \prod_{i>j} (\gamma^{j_{d-i}} - \gamma^{j_{d-j}})$$

can be zero only if some two columns of H are identical. Therefore any d-1 columns of H' and hence any d-1 columns of H are linearly independent. So by theorem 2 the code has minimum distance atleast d.

The lower bound on the minimum distance of a cyclic code obtained in the above manner is called the <u>designed distance</u> of the code.

H Codes : The BCH codes are a class of cyclic codes whose nerator polynomials are chosen to make the minimum distance aranteed by the bound proved in theorem 5.

finition 3 : A cyclic code of length n over GF(q) is a BCH de of designed distance d if for some integer $b \geq 0$.

$$g(X) = \text{LCM} \left[m_b(X), m_{b+1}(X), \ldots, m_{b+d-2}(X) \right],$$

.e., g(X) is the lowest degree monic polynomial having roots $\gamma^b, \gamma^{b+1}, \ldots, \gamma^{b+d-2}$. If b = 1 these are called narrow-sense CH codes. If $n = q^m - 1$, these are called primitive BCH codes for then γ is the primitive element of the field $GF(q^m)$. he pairty check matrix of the above code will be

$$H = \begin{bmatrix} (\gamma^b)^{n-1} & (\gamma^b)^{n-2} & \ldots & \gamma^b & 1 \\ (\gamma^{b+1})^{n-1} & (\gamma^{b+1})^{n-2} & \ldots & \gamma^{b+1} & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ (\gamma^{b+d-2})^{n-1} & (\gamma^{b+d-2})^{n-2} & \ldots & \gamma^{b+d-2} & 1 \end{bmatrix}$$

The most important of the BCH codes are the binary codes obtained by taking γ to be a primitive element of $GF(2^m)$, = 1 and d = 2t+1. These were the codes, whose existence was shown by Bose and Chaudhuri [4]. In this situation f(X) is a code vector if and only if $\gamma, \gamma^2, \ldots, \gamma^{2t}$ are the roots of f(X).

: γ^j is a root of $f(X)$ then $\gamma^{2^m \cdot j}$ is a root for the minimal nction $m_j(X)$ of γ^j for m = 1, 2, 3, ... , and hence a root of (X). Thus f(X) is a codevector if and only if $\gamma, \gamma^3, \ldots, \gamma^{2t-1}$ re roots of f(X) and the generator polynomial of the code is

$$g(X) = \text{LCM}(m_1(X), m_3(X), \ldots, m_{2t-1}(X))$$

he degree of each $m_i(X)$ is less than or equal to m. Therefore (X) has degree atmost mt and the code has atmost mt pairty checks. These results can be summarized in the following theorem which was proved by Bose and Chaudhuri in 1960.

Theorem 6 [4] : For any positive integer m and $t < n/2$, there is a BCH binary code of length $n = 2^m - 1$ which corrects all combinations of t or fewer errors and has no more than mt pairty check symbols.

Nonprimitive BCH codes have received more attention in the second paper of Bose and Chaudhuri [5] . Binary BCH codes are generalized over GF(q) by Gorenstein and Zierler [7].

Example 2 : Two-error-correcting-BCH code over GF(3).

Let

$$GF(3^3) = GF(3)[X] / X^3 + X^2 + 2X + 1$$

as shown in example 1. Let g(X) be the generator polynomial of the BCH code C over GF(3) having roots $\gamma, \gamma^2, \gamma^3, \gamma^4$. Then C has designed distance five and hence by theorem 1 is capable of correcting all error patterns of two or fewer errors and has block

.ength $n = 3^3-1 = 26$. Let $m_1(X)$ be the minimal polynomial for γ^i, $i = 1,2,3,4$. Since the conjugates of γ are γ^3 and γ^9, of γ^2 are γ^6, γ^{18} and of γ^4 are γ^{12}, γ^{10}, degree of $g(X)$ = degree of $m_1(X)$ + degree of $m_2(X)$ + degree of $m_4(X) = 3 + 3 + 3 = 9$. Thus the number of pairty check symbols in the code is 9 and the number of information symbols is $17 = 3^3-1-9$. The pairty check matrix H of the code is

$$H = \begin{bmatrix} 2 & 2 & . & . & . & 0 & 0 \\ 2 & 1 & . & . & . & 1 & 0 \\ 1 & 0 & . & . & . & 0 & 1 \\ 2 & 2 & . & . & . & 1 & 0 \\ 1 & 2 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 0 & 1 \\ 2 & 1 & . & . & . & 2 & 0 \\ 2 & 0 & . & . & . & 1 & 0 \\ 0 & 2 & .. & . & . & 1 & 1 \end{bmatrix}$$

3. THE ENUMERATION OF INFORMATION SYMBOLS IN BCH CODES

In 1963, Henry B. Mann [9] studied the problem of number of information symbols in BCH codes. However, he proved his results under a condition on the designed distance of the codes. Berlekamp studied the same problem in the general form [1]. This chapter summarizes their results which are needed for the work done in Chapter 4.

Let $I(n, d_d)$ denote the number of information symbols in the narrow-sense BCH code C over GF(q) of block length $n = q^m-1$ and designed distance d_d. Let $g(X)$ be the generator polynomial for C over GF(q) whose roots are $\gamma, \gamma^2, \gamma^3, \ldots, \gamma^{d_d-1}$ and their conjugates. If γ^X is a root of the generator polynomial $g(X)$ of the BCH code then $\gamma^{qX}, \gamma^{q^2X}, \gamma^{q^3X}, \ldots$, are conjugates of γ^X. Hence γ^X is a root of $g(X)$, if and only if, there exist some $k(X)$ such that $Xq^k \equiv Y \pmod n$ where $Y < d_d$. Conversely γ^X is a root of the check polynomial $\equiv h(X)$, if and only if, $Xq^k \equiv Y \pmod n$ where $Y \geq d_d$ for all k. Thus we have proved the following result.

Lemma 1 : The number of information symbols in a BCH code over GF(q) of designed distance d_d and block length n is the number of integers X, $1 \leq X \leq n$ satisfying the congruence

$$Xq^k \equiv Y \pmod n \quad , \quad Y \geq d_d \text{ for all } k.$$

This lemma enables us to compute the number of information symbols in the given BCH code over GF(q) without doing any calculation in GF(q) or its extensions. We only have to enumerate certain type of residue classes modulo n. This enumeration is still often tedious in practice, particularly when n and d_d are very large. In order to obtain results for large values of n and d_d, one uses the following alternate form of the above lemma.

Lemma 2 : The number of information symbols in the BCH code over GF(q) of block length n and designed distance d_d is the number of integers i, $0 \leq i \leq n-1$ satisfying the congruence

$$iq^k \equiv j \pmod{n}, \quad 0 \leq j < n+1-d_d, \quad \text{for all } k.$$

Proof : By lemma 1, we know that $I(n, d_d)$ is the number of integers X satisfying

$$Xq^k \equiv y \pmod{n}, \quad 1 \leq X \leq n, \quad y \geq d_d \quad \text{for all } k.$$

or $$nq^k - Xq^k \equiv n-y \pmod{n}, \quad 1 \leq X \leq n, \quad y \geq d_d \quad \text{for all } k.$$

or $$(n-X)q^k \equiv n-y \pmod{n}, \quad 0 \leq n-X \leq n-1, \quad 0 \leq n-y \leq n-d_d \quad \text{for all } k.$$

Replacing nX by i and n-y by j the result follows.

Let $GF(q) = \{0,1,\ldots, q-1\}$ and let $V = V_1V_2\ldots$ be a q-ary sequence, i.e., a sequence whose letters are in GF(q). Then the sequence $\bar{V} = \bar{V}_1 \bar{V}_2 \bar{V}_3 \ldots$, where $\bar{V}_i = (q-1) - V_i$ is called the <u>complement</u> of V. A sequence is said to be a finite

sequence if it has finite number of letters otherwise infinite. If X is a finite sequence and V is any sequence then the concatenation of X and V (denoted by $X * V$) is the sequence $X_1X_2\ldots X_kV_1V_2V_3\ldots$. X is called a prefix and V is called a suffix of the sequence $X*V$. If both X and V are nonempty then X(V) is called a proper prefix (proper suffix) of $X * V$. If X and V are both finite then $V * X$ is a cyclic shift of $X * V$. If X is a finite q-any sequence, say $X = X_1X_2\ldots X_k$, then $\dot{X}$ denotes the sequence $X_1X_2\ldots X_kX_1X_2\ldots X_kX_1\ldots$ In particular, $(Q \pm 1)$ denotes the infinite q-ary sequence $V_1V_2V_3\ldots$ with $V_i = q-1$ for all $i = 1,2,3,\ldots$.

Let X and Y be sequences. $X < Y$ if there exist a positive integer k such that $X_i = Y_i$ for $i = 1,2,\ldots, k-1$ but $X_k < Y_k$. If $X \not< Y$ and $Y \not< X$, then one is a prefix of the other. X is said to an immediate subordinate of Y if X is a finite sequence, $X = X_1X_2\ldots X_k$ with $X_i = Y_i$ for $i = 1,2,\ldots, k-1$ but $X_k < Y_k$. A sequence $Y = Y_1Y_2Y_3\ldots$ has Y_1 immediate subordinates of length one, Y_2 immediate subordinates of length two and in general Y_k immediate subordinates of length k. If the sequence Y has only finite number of nonzero entries, say $Y = Y_1Y_2\ldots Y_k\dot{0}$ then one defines the greatest immediate subordinate of Y as the sequence $Y_1Y_2Y_3\ldots Y_{k-1}(Y_k-1)$. Similarly, Y is an immediate superior of X if there exist some k such that $Y_k > X_k$ and $X_i = Y_i$ for $i=1,\ldots,k-$ If $X = X_1X_2\ldots X_k$ and $X_k \neq q-1$ then the least immediate superior of X is $Y_1Y_2\ldots Y_k$, with $Y_i = X_i$ for $i = 1,\ldots, k-1$ and $Y_k=X_k+$. By definition of immediate subordinates of V, no immediate subordin of V is a proper prefix of any other immediate subordinate of V.

The following lemma was proved by Henry B. Mann in the complementary form.

Lemma 3[9] : Let $J(U,m)$ denotes the number of q-ary m-tuples all of whose cyclic shifts are less than U. If $n = q^m-1$, $n+1-d_d = \sum_{i=1}^{m} U_i q^{m-i}$, $0 \leq U_i < q$ and $U = U_1U_2\ldots U_m$ then

$$I(n, d_d) = J(U, m)$$

Proof: The q-ary m-tuple U corresponds to the integer $n+1-d_d$ Any q-ary m-tuple $W = W_1W_2\ldots W_m$ will correspond to the integer $w = \sum_{i=1}^{m} W_i q^{m-i}$. The first cyclic shift $W_2W_3\ldots W_mW_1$ of W corresponds to

$$\sum_{i=1}^{m-1} W_{i+1} q^{m-i} + W_1 = qw - (q^m-1) W_1 \equiv qw \pmod{n}$$

Similarly it can be shown that the second cyclic shift of W corresponds to q^2w and in general $(m-1)^{th}$ shift corresponds to $q^{m-1}w$. These integers w, qw, ..., $q^{m-1}w$ are less than $n+1-d_d$ if and only if all cyclic shifts of W are less than U. Thus by lemma 2, $I(n, d_d) = J(U, m)$.

In the following theorem by appropriate manipulations on the m-digit q-ary representation of d_d, we derive a simple linear recurrence for a sequence whose m^{th} term is the number of information symbols in the BCH code.

Theorem 4[1] : Let V be a q-ary sequence which exceeds all its proper suffixes. Then

1) Every suffix of every immediate subordinate of V is a concatenation of immediate subordinates of V.

2) Let W be a q-ary sequence. If W and all its cyclic shifts are less than V, then W can be uniquely decomposed into a concatenation of immediate subordinates of V, including a possibly empty end-around immediate subordinate of V, i.e., $W = W^{(1)} * W^{(2)} * \ldots * W^{(j-1)} * W^{(j)}$ where $W^{(1)}, W^{(2)}, \ldots, W^{(j-1)}$ are immediate subordinates of V and $W^{(j)} * W^{(1)} * \ldots * W^{(1)}$ is the end-around immediate subordinate. The end-around immediate subordinate has the prefix $W^{(j)}$ which is the suffix of W and the suffix $W^{(1)} * W^{(2)} * \ldots * W^{(1)}$ which is the prefix of W, as well as concatenation of shorter immediate subordinates $W^{(1)}, W^{(2)}, \ldots, W^{(i)}$.

(3) Every concatenation of immediate subordinates of V including a possibly empty end-around immediate subordinate yields a sequence which has the property that all its cyclic shifts are less than V. No such sequence of length m can exceed the maximum m-digit concatenation of immediate subordinates of V. If Y is the maximum m-digit concatenation of immediate subordinates of V and $Y \leqslant U \leqslant V$ then $J(U,m) = J(V,m)$.

(4) $J(V,m) = m V_m + \sum_{k=1}^{m-1} V_k J(V,m-k)$ where V_j is taken as zero if j exceeds the length of the sequence V.

(5) Let $n = q^m - 1$, $d_d = \sum D_i q^{m-i}$, $0 \leqslant D_i < q$ and let $D = D_1 D_2 \ldots D_m$. If $\bar{V} + (Q-1) < D \leqslant$ least m-digit concatenation of immediate superiors of $\bar{V}$. Then $I(n, d_d) = J(V, m)$.

Proof : (1) Let W be an immediate subordinate of V and let W = P * S where P is the prefix and S is the suffix. Now two cases arise either S is not a proper suffix of W or S is a proper suffix of W. In the first case either S = W in that case S is a concatenation of single immediate subordinate W or S is empty in which case S is a concatenation of empty class of immediate subordinates of V. Now consider the case when S is a proper suffix of W. W = P * S is an immediate subordinate of V implies $V = P * V^{(1)}$. Since V is greater than all its proper suffixes and $W < V$, $S < V^{(1)} < V$. Then S has a prefix say $P^{(1)}$ which is an immediate subordinate of V. Let $S = P^{(1)} * S^{(1)}$. Now $S^{(1)}$ is the suffix of S and S is the suffix of W and so $S^{(1)}$ is the suffix of W. As proved above $S^{(1)}$ can be expressed as $S^{(1)} = P^{(2)} * S^{(2)}$ where $P^{(2)}$ is the immediate subordinate of V. Thus $S = P^{(1)} * P^{(2)} * S^{(2)}$ where $S^{(2)}$ is the suffix of W. Continuing this way we get $S = P^{(1)} * P^{(2)} * \ldots * P^{(k)}$.

(2) Since $W < V$, W has a prefix $W^{(1)}$ which is an immediate subordinate of V, i.e. $W = W^{(1)} * S^{(1)}$. Since $S^{(1)} * W^{(1)}$ is a cyclic shift of W, by hypothesis $S^{(1)} + W^{(1)} < V$. Now two cases arise either $S^{(1)} < V$ or $S^{(1)}$ is a prefix of V. If $S^{(1)} < V$ then $S^{(1)}$ has a prefix $W^{(2)}$ which is an immediate subordinate of V, i.e., $S^{(1)} = W^{(2)} * S^{(2)}$. Hence $W = W^{(1)} + W^{(2)} + S^{(2)}$ where $S^{(2)} < V$ or $S^{(2)}$ is a prefix of V. Proceeding like this we get $W = W^{(1)} * W^{(2)} * \ldots * W^{(j-1)} + W^{(j)}$ where $W^{(j)}$ is a prefix of V and $W^{(1)}, W^{(2)}, \ldots, W^{(j-1)}$ are immediate subordinates of V. Since all cyclic shifts of W are less than U, $W^{(j)} * W^{(1)} * \ldots * W^{(j-}$ has a prefix P which is an immediate subordinate of V. Now $W^{(j)}$ is

prefix of P. If $W^{(j)} * W^{(1)} * \ldots * W^{(i)}$ is a prefix of P and $W^{(j)} * W^{(1)} * \ldots * W^{(i)} * W^{(i+1)}$ is not, let $P = W^{(j)} * W^{(1)} * \ldots * W^{(i)} * S$ where S is a prefix of $W^{(i+1)}$. Since S is the suffix of P, an immediate subordinate of V, by part 1, S is a concatenation of immediate subordinates of V . But no immediate subordinate of V is a proper prefix of any other immediate subordinate of V so S must be empty and hence the result.

(3) Let $W = S^{(j)} * W^{(k)} * W^{(2)} * \ldots * W^{(j-1)} * P^{(j)}$ where $W^{(1)}$, $W^{(2)}$, ..., $W^{(j)}$ are immediate subordinates of V and $W^{(j)} = P^{(j)} * S^{(j)}$ is an end-around immediate subordinate of V. Let C be any cyclic shift of W. Then $C = S^{(k)} * W^{(k+1)} * W^{(k+2)} * \ldots * W^{(j)} * W^{(1)} * \ldots * W^{(k-1)} * P^{(k)}$ where $W^{(k)} = P^{(k)} * S^{(k)}$. Now two cases arise either $S^{(k)}$ is empty or $S^{(k)}$ is not empty. If $S^{(k)}$ is empty then C has a prefix $W^{(k+1)}$ which is an immediate subordinate of V and hence $C < V$. If $S^{(k)}$ is not empty then $S^{(k)}$ is the suffix of the immediate subordinate $W^{(k)}$ of V and hence $S^{(k)}$ can be expressed as a concatenation of immediate subordinates of V by part 1. Therefore C has a prefix which is an immediate subordinate of V and hence $C < V$. Clearly no such sequence of length m can exceed the maximum m-digit concatenation of immediate subordinates of V. If Y is the maximum m-digit concatenation of immediate subordinates of V and $Y \leqslant U \leqslant V$ then $J(U,m) = J(V,m)$.

(4) If $V = V_1 V_2 \ldots V_m$, then V has V_m immediate subordinates of length m, each of which has m distinct cyclic shifts. Thus W may chosen as a single end-around immediate subordinate of V in mV_m ways.

If W is the concatenation of several immediate subordinates of V, i.e., $W = W^{(1)} * W^{(2)} * \ldots * W^{(j-1)} + W^{(j)}$ where $W^{(1)}, W^{(2)}, \ldots, W^{(j-1)}$ are immediate subordinates of V and $W^{(j)}$ is the proper prefix of the immediate subordinate $W^{(j)} * W^{(1)} * \ldots * W^{(i)}$. Then the length of W is the length of $W^{(j-1)}$ plus the length of $W^{(1)} * \ldots * W^{(j-2)} * W^{(j)}$. For each k there are V_k choices of $W^{(j-1)}$ of length k and J(V, m-k) choices for $W^{(1)} * \ldots * W^{(j-2)} * W^{(j)}$. Therefore

$$J(V, m) = m V_m + \sum_{k=1}^{m-1} V_k J(V, m-k).$$

(5) Consider J(U, m). Since J(U, m) is a monotonic function of U, it is sufficient to prove the result for the two extreme cases. Considering the least special case. Let D be the least m-digit q-ary sequence greater than $\bar{V} + (Q \pm 1)$. Letting $d_d = \sum_i D_i q^{m-i}$, $v = \sum_i V_i q^{m-i}$, $\bar{v} = \sum_i \bar{V}_i q^{m-i}$.

$$\begin{aligned} d_d + v &= \sum_i D_i q^{m-i} + \sum_i V_i q^{m-i} \\ &= \sum_i \bar{V}_i q^{m-i} + 1 + \sum_i V_i q^{m-i} \\ &= \sum_i (\bar{V}_i + V_i) q^{m-i} + 1 \\ &= q^m - 1 + 1 = q^m = n+1 \end{aligned}$$

or $v = n + 1 - d_d$.

Hence by lemma 3 $I(n, d_d) = J(V, m)$.

Considering the greatest special case. Let $\bar{D}$ be the least m-digit concatenation of immediate superior of $\bar{V}$. Then $\bar{D}$ is the greatest m-digit concatenation of immediate inferiorates of V. In notation of part 3, $\bar{D} = Y$. Letting $\bar{\delta}_d = \sum_i \bar{D}_i q^{m-i}$, we have $\bar{\delta}_d = n-d_c$. Then by lemma 3 and part 3 above,

$$I(n,d_d) = J(U,m) = J(V,m)$$

The above theorem will enables us to determine the number of information symbols in q-ary BCH code of block length $n = q^m-1$ and designed distance $\delta_d = \sum_i D_i q^{m-i}$ if we can find a sequence V which is greater than all its proper suffixes and has the property :

$\bar{V}+(\lambda-1) < D \leq$ least m-digit concatenation of immediate superiors of $\bar{V}$. In the following theorem the problem of finding such a V is reduced to the problem of finding λ, which is a prefix of $\bar{D}$.

Theorem 2.1 : Let X be the shortest prefix of $\bar{D}$ such that $\bar{D} = X*F$, $F+(\lambda-1) \geq \bar{D}+(\lambda-1)$ and let V be the least immediate superior of X. Then

(1) $\bar{V}+(\lambda-1) < D \leq$ least m-digit concatenation of immediate superiors of $\bar{V}$.

(2) V exceeds all its proper suffixes.

Proof :(1) Since X is a prefix of $\bar{D}$ and V is an immediate superior of X, V is an immediate superior of $\bar{D}$. So $V > \bar{D}$ and $V > \bar{D}+(\lambda-1)$.

complementing this we get $\bar{V}*(\overset{\bullet}{Q}-1) < L*\overset{\bullet}{0}$ so $\bar{V}*(\overset{\bullet}{Q}-1) < D$.

Let $X^{(k)} = X*X*\ldots*X$ ($\leftarrow$ k times). Then $F*(\overset{\bullet}{Q}-1) \geq \bar{D}*(\overset{\bullet}{Q}-1)$ is equivalent to $V*(\overset{\bullet}{Q}-1) \geq X^{(1)}*F*(\overset{\bullet}{Q}-1)$. Therefore, $X*F*(\overset{\bullet}{Q}-1) \geq X*X^{(1)}*F*(\overset{\bullet}{Q}-1)$, or $X*F*(\overset{\bullet}{Q}-1) \geq X^{(2)}*F*(\overset{\bullet}{Q}-1)$. By induction, $X^{(k)}*F*(\overset{\bullet}{Q}-1) \geq X^{(k+1)}*F*(\overset{\bullet}{Q}-1)$, and $\bar{D}*(\overset{\bullet}{Q}-1) \geq X^{(k)}*F*(\overset{\bullet}{Q}-1)$ for all k. Since this is true for arbitrary large k, $\bar{D}*(\overset{\bullet}{Q}-1) \geq \overset{\bullet}{X}$. Complementing it we get $L*\overset{\bullet}{0} \leq \overset{\bullet}{X}$ $\leq$ any infinite concatenation of immediate superiors of $\bar{V}$. Therefore, $D \leq$ any m-digit concatenation of immediate superiors of $\bar{V}$.

(2) Let $X = Y*Z+L$, where Y and Z are arbitrary and L is the final digit of X. We have $V = Y*Z*(L+1)$, $\bar{D} = Y*Z*L*F$. $F*(\overset{\bullet}{Q}-1) \geq Y*Z*L*F*(\overset{\bullet}{Q}-1) = X+1*(\overset{\bullet}{Q}-1)$. If $Y*Z*L*F*(\overset{\bullet}{Q}-1) \leq Z*L*F*(\overset{\bullet}{Q}-1)$ then Y is a shorter prefix than X satisfying the given condition. Thus $Y*Z*L*F*(\overset{\bullet}{Q}-1) > Z*L*F*(\overset{\bullet}{Q}-1)$. If some proper suffix of V say $Z*(L+1)$ exceeds V then $Z*(L+1) > Y*Z*(L+1) > Y*Z*L = X$. If $Z*L > X$ then $Z*L*F*(\overset{\bullet}{Q}-1) > X*F*(\overset{\bullet}{Q}-1)$, a contradiction. On the other hand if $*L$ is a prefix of X then $X = Z+L+G$ and from $X*F*(\overset{\bullet}{Q}-1) > Z*L*F(\overset{\bullet}{Q}-1)$ we have $Z*L*G*F*(\overset{\bullet}{Q}-1) > Z*L*F*(\overset{\bullet}{Q}-1)$, and $G*F*(\overset{\bullet}{Q}-1) > F*(\overset{\bullet}{Q}-1) \geq X*F*(\overset{\bullet}{Q}-1)$. Since $*L$ is a shorter prefix than X, this leads to a contradiction. Therefore $Z*(L+1) < Y*Z*(L+1)$, i.e., V precedes all its own proper suffixes.

Example : Number of information symbols in two-error correcting BCH code over GF(3). Considering the example 2 of chapter 2 with $q = 3$, $n = 26$, $d_0 = 5$, $D = 012$, $\bar{D} = 210$, $\bar{D}*\overset{\bullet}{2} = 210222\ldots$, $X=210$, $V=211$.

m	1	2	3
J(V,m)	2	6	17

Thus the number of information symbols in the code is 17.

Now in the following section asymptotic results for $I(n,d_d)$ as n and d_d both approach infinity with their ratio remaining fixed have been obtained. Define the enumerator

$$J(V,z) = \sum_{m=1}^{\infty} J(V,m)\, z^m$$

Given a sequence V which exceeds all its proper suffixes, we may also define

$$V(z) = \sum_{k} V_k z^k .$$

So that

$$zV'(z) = \sum_{k} k\, V_k\, z^k$$

From theorem 1

$$J(V,m) = m\, V_m + \sum_{k=1}^{m-1} V_k J(V,m-k) ,$$

$$\text{or } \sum_{m=1}^{\infty} J(V,m)z^m = \sum_{m=1}^{\infty} m\, V_m\, z^m + \sum_{m=1}^{\infty} \sum_{k=1}^{m-1} V_k J(V,m-k)z^m,$$

$$\text{or } J(V,z) = zV'(z) + \left(\sum_{m=1} J(V,m-k)\, z^{m-k} \right) \left(\sum_{k=1}^{m-1} V_k\, z^k \right) ,$$

$$= zV'(z) + V(z)J(V,z),$$

$$\text{or} \qquad J(V,z) = \frac{zV'(z)}{1-V(z)} .$$

Let $\rho_1, \rho_2, \ldots$ be the complex reciprocal roots of $1-V(z)$. Then

$$1-V(z) = \prod_{i} (1-\rho_i z)$$

$$\text{or} \qquad -V'(z) = -\sum_i \rho_i \prod_{j\neq i} (1-\rho_j z)$$

$$\text{or} \qquad zV'(z) = \sum_i \rho_i z \prod_{j\neq i} (1-\rho_j z)$$

$$\text{or} \qquad \frac{zV'(z)}{1-V(z)} = \sum_i \frac{\rho_i z}{1-\rho_i z} = \sum_i \sum_{m=1}^{\infty} (\rho_i z)^m$$

$$= \sum_i \sum_{m=1}^{\infty} \rho_i^m z^m$$

$$\text{Therefore } J(V,z) = \sum_{m=1}^{\infty} \sum_i \rho_i^m z^m$$

$$\text{or} \qquad J(V,m) = \sum_i \rho_i^m$$

This proves the following theorem.

Theorem 3[1] : $J(V,m) = \sum_i \rho_i^m$ where ρ_i's are complex numbers defined by the equation

$$1-V(z) = \prod_i (1-\rho_i z) .$$

Although the above theorem gives an explicit expression for $J(V,m)$, the expression depends upon the complex numbers ρ_i. For finite values of m, it is usually easier to compute $J(V,m)$ directly from the recurrence relation of theorem 1, since these calculations involve only integers. For asymptotic results, however the above equation is very useful.

Let $\rho = \max_i \{|\rho_i|\}$

and let

$$s = \log_q \rho$$

Since the coefficients of $V(z)$ are nonnegative integers not exceeding $q-1$, $1 \leq \rho < q$. Clearly $J(V,m)$ is asymptotically equal to ρ^m and we write

$$J(V,m) \sim \rho^m$$

for large m in the sense that

$$\lim_{m\to\infty} \rho^{-m} J(V,m) = 1$$

Similarly,

$$\log_q J(V,m) \sim \log_q \rho^m = m \log_q \rho = ms$$

or $$J(V,m) \sim q^{ms}$$

If $u = \sum_{i=1}^{\infty} U_i q^{-i}$, $U = U_1 U_2 U_3 \ldots$ and $\dot{X} \leq \overline{U} \leq V + \dot{O}$

where V exceeds all its proper suffixes and X is the maximum immediate subordinate of V, then

$$I(q^m-1, uq^m) \sim q^{ms}$$

In other words, if we fix the fraction $d_d/n = u$, n and d_d grow large then

$$I(q^m-1, uq^m) \sim n^{s(u)}$$

where

$$s(u) = \lim_{m\to\infty} \frac{\log_q I(q^m-1, uq^m)}{m}$$

For a given q, the function s(u) is rather complicated. However we can compute s(u) in the following way. Let $0 < u < 1$. Then $u = \sum_{i=1}^{\infty} U_i q^{-i}$. Let $U = U_1 U_2 U_3 \ldots$ and let $\bar{U}$ be the complement of U. If $\bar{U}$ exceeds all its proper suffixes let $V = \bar{U}$, otherwise let V be the least immediate superior of X where X is the shortest prefix with $\bar{U} = X*P$ and $\bar{U} \leqslant P$. Let $V(z) = \sum_i V_i z^i$ where $V = V_1 V_2 \ldots$ and let $1-V(z) = \prod_i (1-\rho_i z)$. If $\rho = \max \{|\rho_i|\}$ then

$$s(u) = \log_q \rho \qquad \ldots(1)$$

4. [illegible]

In [2] Berlekamp has proved that long primitive binary BCH codes have distance $d \sim \frac{2n\ln k^{-1}}{\log_2 n}$. In this chapter this result has been generalized to the BCH codes over GF(q). The results of this chapter are based on chapter 3. Following the notations of chapter 3, let $I(n,d_d)$ denote the number of information symbols in the BCH code over GF(q) of block length n and designed distance d_d. In chapter 3 we have seen that the number of information symbols in long primitive BCH codes of fixed ratio $d_d/n = u$ is asymptotic to $n^{s(u)}$ in the sense that

$$\lim_{m\to\infty} \frac{I(q^m-1, uq^m)}{n^{s(u)}} = 1 .$$

We first investigate the functional relationship between u and s when $u \downarrow 0$ (u approches 0 from right) and $s \uparrow 1$ (s approaches 1 from left). It turns out to be slightly easier to investigate another function r(u) defined in the following manner. Let $0 < u < 1$. Then

$$u = \sum_{i=1}^{\infty} U_i q^{-i}, \quad U_i \in \{0,1,\ldots, q-1\} \text{ for all } i = 1,2,3,\ldots \quad (1)$$

In this chapter a summation unless specified will represent a summation from 1 to ∞. Corresponding to the sequence $U_1U_2U_3\ldots$ we associate a complementary sequence $V_1V_2V_3\ldots$ with $V_i = q-1-U_i$ for all $i = 1,2,3,\ldots$ Define the corresponding analytic functions

$$U(z) = \sum_i U_i z^i \quad \ldots (2)$$

and $$V(z) = \sum_i V_i z^i$$

With this notation equation (1) may be restated as

$$u = U(1/q) \qquad \ldots(3)$$

For positive real z, function V(z) is monotonically increasing. Since V(0) = 0 and V(1) ≥ 1, the equation V(z) = 1 has a unique positive real root, say X. Now

$$U(z) + V(z) = \sum_{i=1}^{\infty} U_i z^i + \sum_{i=1}^{\infty} V_i z^i = \sum_{i=1}^{\infty} (q-1)z^i = \frac{(q-1)z}{1-z}$$

implies that X is also the unique positive real root of the equation

$$U(X) = \frac{(q-1)X}{1-X} - 1 = \frac{qX-1}{1-X} \qquad \ldots (4)$$

Evidently $X \geq \frac{1}{q}$. Define

$$r(u) = -\log_q X \qquad \ldots (5)$$

Here in this chapter ln represents logarithm to the base e and log represents logarithm to the base q. For the function U(z) defined by (2)

$$U'(z) = \sum_i i U_i z^{i-1}, \quad zU'(z) = \sum_i i U_i z^i,$$

$$zU'(z)\ln z = \sum_i i U_i z^i \ln z, \quad zU'(z)\ln z = \sum_i U_i z^i \ln z^i,$$

and $$\frac{zU'(z)\ln z}{U(z)} = \frac{\sum U_i z^i \ln z^i}{U(z)}$$

$$= \sum U_i \frac{z^i}{U(z)} \ln \left(\frac{z^i}{U(z)}\right) + \ln U(z)$$

$$= -H(z) + \ln U(z) \qquad \ldots (6)$$

where $H(z) = -\sum U_i \frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right)$ ... (7)

Now we shall determine bounds on $H(z)$ and $H'(z)$. Differentiating (7) we get

$$-H'(z) = \sum U_i \left[\frac{iz^{i-1}}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) - \frac{z^i}{U^2(z)} U'(z) \ln\left(\frac{z^i}{U(z)}\right) + \left(\frac{iz^i}{U(z)} - \frac{z^i}{U^2(z)} U'(z)\right) \right]$$

$$= \sum U_i \left(\frac{iz^{i-1}}{U(z)} - \frac{z^i}{U^2(z)} U'(z) \right) \left(1 + \ln\left(\frac{z^i}{U(z)}\right)\right)$$

Hence using (6),

$$H'(z)\, z \ln(1/z) = \sum U_i \left(\frac{iz^i \ln z}{U(z)} - \frac{z^i}{U(z)} \cdot \frac{z \ln z\, U'(z)}{U(z)} \right)\left(1 + \ln\left(\frac{z^i}{U(z)}\right)\right)$$

$$= \sum U_i \left(\frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) + \frac{z^i}{U(z)} H(z)\right) \left(1 + \ln\left(\frac{z^i}{U(z)}\right)\right),$$

$$= -H(z) + H(z) + \sum U_i \frac{z^i}{U(z)} \ln^2\left(\frac{z^i}{U(z)}\right) - H^2(z).$$

Now $z \ln 1/z$, $\sum U_i \frac{z^i}{U(z)} \ln^2(z^i/U(z))$ and H^2 are all nonnegative for $0 < z < 1$ implies

$$-H^2 \leq H'(z)\, z \ln(1/z) \leq \sum U_i \frac{z^i}{U(z)} \ln^2\left(\frac{z^i}{U(z)}\right) \quad \text{... (8)}$$

Define the sequences $U^* = U_1^* U_2^* U_3^* \ldots$ and $U^{**} = U_1^{**} U_2^{**} U_3^{**}$ by

$$U_i^* = \begin{cases} 1 & \text{if } U_i \neq 0 \\ 0 & \text{otherwise} \end{cases}, \quad U_i^{**} = \begin{cases} q-1 & \text{if } U_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

and the corresponding analytic functions

$$U^*(z) = \sum_i U_i^* z^i \quad , \quad U^{**}(z) = \sum_i U_i^{**} z^i$$

Note that $U^{**}(z) = (q-1)\, U^*(z)$. If there are j or more ones in U^*, let $i(j)$ denote the index of the j^{th} one and if there are fewer than j ones take $i(j) = \infty$ and let $z^\infty = 0$. Let

$$p_j = \frac{z^{i(j)}}{U^*(z)} \qquad \ldots (9)$$

Since $z^i/U(z) \geq z^i/U^{**}(z)$ and ln X is an increasing function,

$$\ln\left(\frac{z^i}{U(z)}\right) \geq \ln\left(\frac{z^i}{U^{**}(z)}\right) \qquad \ldots (10)$$

Moreover,

$$U_i \frac{z^i}{U(z)} \leq U_i^{**} \frac{z^i}{U(z)} \leq (q-1)\, U_i^{**} \frac{z^i}{(q-1)\, U^*(z)}$$

Or $$U_i \frac{z^i}{U(z)} \leq (q-1)\, U_i^{**} \frac{z^i}{U(z)} \qquad \ldots (11)$$

Using (10) and (11) we get

$$U_i \frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) \geq (q-1)\, U_i^{**} \frac{z^i}{U^{**}(z)} \ln\left(\frac{z^i}{U^{**}(z)}\right) \text{ for all } i=1,2,\ldots$$

and so

$$\sum_i U_i \frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) \geq \sum_i (q-1)\, U_i^{**} \frac{z^i}{U^{**}(z)} \ln\left(\frac{z^i}{U^{**}(z)}\right),$$

or

$$-\sum_i U_i \frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) \leq -\sum_i (q-1)\, U_i^{**} \frac{z^i}{U^{**}(z)} \ln\left(\frac{z^i}{U^{**}(z)}\right),$$

or $$H(z) \leq (q-1)\, H^{**}(z) \qquad \ldots (12)$$

where $$H^{**}(z) = -\sum U_i^{**} \frac{z^i}{U^{**}(z)} \ln\left(\frac{z^i}{U^{**}(z)}\right),$$

$$= -\sum U_i^{*+} \frac{z^i}{(q-1)\, U^*(z)} \ln\left(\frac{z^i}{(q-1)U^*(z)}\right),$$

$$= -\sum_{i,\, U_i \neq 0} \frac{z^i}{U^*(z)} \ln\left(\frac{z^i}{U^*(z)}\right) + \ln(q-1) \sum_{i, U_i \neq 0} \frac{z^i}{U^*}$$

$$= H^*(z) + \ln(q-1) \sum U_i^* \frac{z^i}{U^+(z)},$$

or

$$H^{**}(z) = H^*(z) + \ln(q-1) \qquad \ldots (13)$$

where $$H^+(z) = -\sum \frac{z^i}{U^*(z)} \ln\left(\frac{z^i}{U^*(z)}\right).$$

Using (9) we get

$$H^*(z) = -\sum_{j=1}^{\infty} p_j \ln p_j ,$$

$$= -p_1 \ln p_1 - (1-p_1) \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln\left(\frac{p_j}{1-p_1}\right) - (1-p_1) \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln(1-p_1$$

$$= -p_1 \ln p_1 - (1-p_1)\left[\ln(1-p_1) + \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln\left(\frac{p_j}{1-p_1}\right)\right].$$

$$= -p_1 \ln p_1 - (1-p_1)\left[\ln(1-p_1) - H_2^*(z) \right].$$

where $H_2^*(z) = -\sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln\left(\frac{p_j}{1-p_1}\right)$.

Both H^* and H_2^* are entropies of the form : $\Sigma\, q_j \ln q_j$ where $0 \leq q_{j+1} \leq z\, q_j$ for all j and $\Sigma\, q_j = 1$. It follows that H^* and H_2^* have a common upper bound H_{up}^* which satisfies

$$H_{up}^*(z) = -p_1 \ln p_1 - (1-p_1)\left[\ln(1-p_1) - H_{up}^*(z) \right].$$

Hence on simplifying we have,

$$H_{up}^*(z) = -\ln p_1 - \frac{1-p_1}{p_1} \ln(1-p_1).$$

Differentiating with respect to p_1 we get

$$\frac{d H_{up}^*}{dp_1} = -\frac{1}{p_1} + \frac{\ln(1-p_1)}{p_1} + \frac{1}{p_1} + \frac{(1-p_1)\ln(1-p_1)}{p_1^2}$$

$$= \frac{\ln(1-p_1)}{p_1^2} < 0 \qquad \text{for } 0 < p_1 < 1 .$$

Moreover $d H^*_{up}/dp_1 = 0$ implies $p_1 = 0$. Therefore H^*_{up} is maximized by taking p_1 as small as possible. But since $q_{j+1} \leq z\, q_j \leq z^j q_1$, it follows that $\sum_j q_j \leq q_1 \sum_j z^{j-1}$, $q_1 \geq 1-z$ and

$$H^*_{up}(z) \leq -\ln(1-z) - \frac{z}{1-z} \ln z ,$$

$$= \frac{-(1-z)\ln(1-z) - z \ln z}{1 - z} .$$

Consider the function $f(z) = -(1-z)\ln(1-z) - z \ln z$. Then $f'(z) = 1 + \ln(1-z) - 1 - \ln z = 0$ implies that the maximum value of $f(z)$ is attained at $z = 1/2$ and the maximum value is $-1/2 \ln 1/2 - 1/2 \ln 1/2 = 1/2 \ln 2 + 1/2 \ln 2 = \ln 2$. Therefore,

$$H^*(z) \leq H^*_{up}(z) \leq \frac{\ln 2}{1-z} \qquad \ldots (14)$$

Substituting the value in (13) and (12) and noting that $H(z)$ is bounded below by 0 we get,

$$H^{**}(z) \leq \frac{\ln 2}{1-z} + \ln(q-1)$$

and

$$0 \leq H(z) \leq (q-1)\left(\frac{\ln 2}{1-z} + \ln(q-1)\right) \qquad \ldots (15)$$

Consider the function

$$D_1(z) = \sum_i U_i \frac{z^i}{U(z)} \ln^2\left(\frac{z^i}{U(z)}\right) .$$

As $\ln^2 x$ is a decreasing function in the open interval (0,1),

$$\ln^2\left(\frac{z^i}{U(z)}\right) \leq \ln^2\left(\frac{z^i}{U^{**}(z)}\right).$$

Hence using (11) we get

$$U_i \frac{z^i}{U(z)} \ln^2\left(\frac{z^i}{U(z)}\right) \leq (q-1)\, U_i^{**} \frac{z^i}{U^{**}(z)} \ln^2\left(\frac{z^i}{U^{**}(z)}\right) \text{ for } i = 1,2,\ldots$$

So

$$\sum_i U_i \frac{z^i}{U(z)} \ln^2\left(\frac{z^i}{U(z)}\right) \leq (q-1) \sum_i U_i^{**} \frac{z^i}{U^{**}(z)} \ln^2\left(\frac{z^i}{U^{**}(z)}\right),$$

or $$D_1 \leq (q-1) \sum_{i,U_i\neq 0} (q-1) \frac{z^i}{(q-1)U^*(z)} \ln^2\left(\frac{z^i}{(q-1)\,U^*(z)}\right),$$

$$= (q-1) \sum_i \frac{z^i}{U^*(z)} \left(\ln\left(\frac{z^i}{U^*(z)}\right) - \ln(q-1)\right)^2,$$

$$= (q-1) \sum_i \frac{z^i}{U^*(z)} \ln^2\left(\frac{z^i}{U^*(z)}\right) + (q-1)\ln^2(q-1) + 2(q-1)\ln(q-1)H^*$$

Let $D_1^*(z) = \sum_i U_i^* \frac{z^i}{U^*(z)} \ln^2\left(\frac{z^i}{U^*(z)}\right)$. Then using (14) we have

$$D_1 \leq (q-1)\, D_1^* + (q-1)\ln^2(q-1) + 2(q-1)\frac{\ln(q-1)\ln 2}{1-z} \quad \ldots(16)$$

Using (9) one gets $D_1^* = \sum_j p_j \ln^2 p_j$. Let

$$D_2^* = \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln^2\left(\frac{p_j}{1-p_1}\right) \qquad \ldots (17)$$

Then

$$D_2^* = \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln^2 p_j + \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln^2(1-p_1)$$

$$- 2 \ln(1-p_1) \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln p_j ,$$

$$= \frac{D_1^* - p_1 \ln^2 p_1}{1-p_1} - \ln^2(1-p_1) - 2 \ln(1-p_1) \sum_{j=2}^{\infty} \frac{p_j}{1-p_1} \ln \frac{p_j}{1-p_1} ,$$

$$= \frac{D_1^* - p_1 \ln^2 p_1}{1-p_1} - \ln^2(1-p_1) + 2 \ln(1-p_1) \; H_2^*(z) .$$

Hence

$$D_1^* - (1-p_1) D_2^* = p_1 \ln^2 p_1 + (1-p_1) \ln^2(1-p_1) - 2 (1-p_1) \ln(1-p_1) H_2^*(z$$

Using (14) we get,

$$D_1^* - (1-p_1) D_2^* \leqslant p_1 \ln^2 p_1 + (1-p_1) \ln^2(1-p_1) - 2(1-p_1) \ln(1-p_1) \frac{\ln 2}{1-z} .$$

Since D_1^* and D_2^* have the same form, they have a common upper bound D_{up}^* which satisfies

$$D_{up}^* \leqslant \ln^2 p_1 + \frac{(1-p_1)}{p_1} \ln^2(1-p_1) - \frac{2(1-p_1) \ln(1-p_1)}{p_1} \cdot \frac{\ln 2}{1-z} .$$

But since $\ln^2 p$, $\frac{1-p}{p} \ln^2(1-p)$ and $- \frac{(1-p) \ln(1-p)}{p}$ are monotonical decreasing functions of p, it follows that D_{up}^* is maximized by

taking p_1 as small as possible. Setting $p_1 = 1-z$ gives

$$D^*_{up} = D^*_{up}(z) \leq \ln^2(1-z) + \frac{z}{1-z}\ln^2 z - \frac{2z \ln z}{(1-z)^2}.\ln 2$$

Substituting the above value in (16) we have

$$D_1(z) \leq (q-1)\left[\ln^2(1-z) + \frac{z}{1-z}\ln^2 z - \frac{2z \ln z}{(1-z)^2}\ln 2\right]$$

$$+ (q-1)\ln^2(q-1) + \frac{2(q-1)\ln(q-1)\ln 2}{1-z}.$$

Substituting the value of $D_1(z)$ in (8) we get

$$H'(z)z\ln(1/z) \leq (q-1)\left[\ln^2(1-z) + \frac{z}{1-z}\ln^2 z - \frac{2z \ln z}{(1-z)^2}.\ln 2\right]$$

$$+ (q-1)\ln^2(q-1) + \frac{2(q-1)\ln(q-1)\ln 2}{1-z} \quad \ldots(1$$

Also from (8) and (15) we get

$$-(q-1)^2\left(\frac{\ln 2}{1-z} + \ln(q-1)\right)^2 \leq H'(z)\, z \ln(1/z) \quad \ldots (19)$$

Equations (18) and (19) give upper and lower bounds on $H'(z)$.

On putting $z = 1/q$ in (18) and (19), we get

$$-\frac{q(q\ln 2 + (q-1)\ln(q-1))^2}{\ln q} \leq H'(1/q) \leq \frac{q(q-1)}{\ln q}\left[\ln^2\frac{q-1}{q} + \frac{1}{q-1}\ln^2\right.$$

$$\left. - \frac{2q \ln(1/q)\ln 2}{(q-1)^2} + (q-1)\ln^2(q-1) + \frac{2q \ln(q-1).\ln 2}{1}\right] \ldots (20)$$

Thus $H'(1/q)$ is bounded.

The next theorem gives an estimate on the size of u for which r(u) attains any given value near 1.

Theorem 1: If $r(u) = 1 - \frac{\epsilon}{\log_e q^q}$ ...(21)

and ϵ is small then

$$u = \left(\frac{\epsilon}{q-1}\right)^{1+\frac{\epsilon}{\ln q^q}}\left(1+\frac{\epsilon}{q-1}\left(\frac{q+1}{2q}-\frac{\ln\ln U(1/q)}{\ln q^q}\right)+O\left(\epsilon^2 \ln\frac{\epsilon}{q-1}\right)\right).$$

Proof. We have seen earlier that if $U(z) = \sum U_i z^i$ then

$$\frac{U'(z)z\ln z}{U(z)} = \ln U(z) + \sum U_i \frac{z^i}{U(z)} \ln\left(\frac{z^i}{U(z)}\right) = \ln U(z) - H(z)$$

Dividing both sides by $\ln U(z)$ we have

$$1 - \frac{H(z)}{\ln U(z)} = \frac{U'(z)z\ln z}{U(z)\ln U(z)} = \frac{(\ln\ln U(z))^{-1}}{(\ln\ln z^{-1})^{-1}} \qquad \ldots(22)$$

Let X be defined by the equation (4). Since $\frac{1}{q} \leq X$, the chain rule for the derivatives and the mean value theorem guarantee the existence of a y, $\frac{1}{q} \leq y \leq X$ such that

$$\frac{\ln\ln U(X)^{-1} - \ln\ln (1/q)^{-1}}{\ln\ln X^{-1} - \ln\ln q} = 1 - \frac{H(y)}{\ln U(y)} \qquad \ldots(23)$$

From equations (8) and (21)

$$\ln\ln X^{-1} - \ln\ln q = \ln\left(-\frac{\ln X}{\ln q}\right) = \ln r(u) = \ln\left(1 - \frac{\epsilon}{\ln q^q}\right)$$

$$= -\frac{\epsilon}{\ln q^q} + O(\epsilon^2)$$

so that (23) becomes

$$\left(-\frac{\epsilon}{\ln q^q}+O(\epsilon^2)\right)\left(1-\frac{H(y)}{\ln U(y)}\right)=\ln U(x)^{-1}-\ln\ln u^{-1}$$

$$=\ln\left(\frac{\ln U(x)}{\ln\frac{\epsilon}{q-1}}\right)-\ln\left(\frac{\ln u}{\ln\frac{\epsilon}{q-1}}\right) \qquad \ldots(24)$$

From (5) and (21) we get

$$u=q^{-r}=\frac{1}{q}e^{\epsilon/q}=\frac{1}{q}\left[1+\frac{\epsilon}{q}+\frac{\epsilon^2}{2q^2}+O(\epsilon^3)\right]$$

Therefore

$$U(x)=\frac{1\left[1+\frac{\epsilon}{q}+\frac{\epsilon^2}{2q^2}+O(\epsilon^3)\right]-1}{1-\frac{1}{q}\left[1+\frac{\epsilon}{q}+O(\epsilon^2)\right]} \qquad \ldots(25)$$

$$=\frac{\epsilon}{q-1}\left(1+\frac{(q+1)\epsilon}{2q(q-1)}+O(\epsilon^2)\right) \qquad \ldots(25)$$

and hence

$$\ln U(x)=\ln\frac{\epsilon}{q-1}+\frac{(q+1)\epsilon}{2q(q-1)}+O(\epsilon^2)$$

$$\frac{\ln U(x)}{\ln\frac{\epsilon}{q-1}}=1+\frac{(q+1)\epsilon}{2(q-1)q\ln\frac{\epsilon}{q-1}}+O\left(\frac{2}{\ln\frac{\epsilon}{q-1}}\right).$$

Combining this with (24) gives

$$\ln\left(\frac{\ln u}{\ln\frac{\epsilon}{q-1}}\right)=\frac{\epsilon}{\ln q^q}+\frac{(q+1)\epsilon}{2(q-1)q\ln\frac{\epsilon}{q-1}}-\frac{\epsilon H(y)}{\ln q^q\ln U(y)}$$

$$+O(\epsilon^2)+\frac{H(y)}{\ln U(y)}O(\epsilon^2) \qquad \ldots(26)$$

Since $y \leq 1$ y (2), we have

$$U(y) \leq U(1) = \frac{\epsilon}{q-1} + O(\epsilon^2) .$$

$$\frac{1}{\ln U(y)} = O\left(\frac{1}{\ln \frac{\epsilon}{q-1}}\right)$$

From equation (1), we have, $H(y) = O(1)$ so

$$\ln\left(\frac{\ln u}{\ln \frac{\epsilon}{q-1}}\right) = O\left(\frac{\epsilon}{q-1}\right) ,$$

$$u = \frac{\epsilon}{q-1} + O\left(\frac{\epsilon^2}{(q-1)^2} \ln \frac{\epsilon}{q-1}\right) \qquad \ldots (27)$$

Since $\frac{1}{q} \leq y \leq 1$, $U(1/q) \leq U(y) \leq U(1)$, (2) and (27) gives

$$U(y) = \frac{\epsilon}{q-1}\left(1 + O\left(\frac{\epsilon}{q-1} \ln \frac{\epsilon}{q-1}\right)\right)$$

Taking ln both sides we get

$$\ln U(y) = \ln \frac{\epsilon}{q-1} + O\left(\frac{\epsilon}{q-1} \ln \frac{\epsilon}{q-1}\right)$$

$$= \ln \frac{\epsilon}{q-1}\left(1+O\left(\frac{\epsilon}{q-1}\right)\right) \qquad \ldots (28)$$

Hence from equation (20) we have, that $H(z)$ is differentiable at $z = 1/q$ and $H'(1/q) = O(1)$ and so

$$H(y) = H(1/q) + O(\epsilon/q-1).$$

Combining this with (26) and (28) gives

$$\ln\left(\frac{\ln u}{\ln \frac{\epsilon}{q-1}}\right) = \frac{\epsilon}{\ln q^q} + \frac{(q+1)\epsilon}{2(q-1)q \ln \frac{\epsilon}{q-1}} - \frac{\epsilon H(1/q)}{\ln q^q \ln \frac{\epsilon}{q-1}} + O(\epsilon^2)$$

$$u = \left(\frac{\epsilon}{q-1}\right)^{1+\frac{\epsilon}{\ln q^q}} \cdot \left(1+\frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)H(1/q)}{\ln q^q}\right) + O(\epsilon^2 \ln \frac{\epsilon}{q-1})\right)$$

where $H(1/q)$ is a function of u. Let $H(1/q) = h(u)$. Then

$$u = \left(\frac{\epsilon}{q-1}\right)^{1+\frac{\epsilon}{\ln q^q}} \cdot \left(1 + \frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)h(u)}{\ln q^q}\right) + O(\epsilon^2 \ln \frac{\epsilon}{q-1})\right)$$

or $r = 2$, u is approximated by

$$u = \epsilon^{1+\epsilon/\ln 4} \cdot \left(1 + \epsilon\left(\frac{3}{4} - \frac{h(u)}{\ln 4}\right) + O(\epsilon^2 \ln \epsilon)\right)$$

where $r(u) = 1-\epsilon/\ln 4$ and ϵ is small. This result was proved by [illegible] 2]. In the following theorem we will determine the values of u for which r(u) and s(u) are equal.

Theorem 2 : For a given u, $0 < u < 1/q$, there exist numbers t and w such that $t \leq u \leq w$ and $s(u) = r(t) = r(w)$ where functions s(u) and r(u) are defined by (1) in chapter 3 and (*) in the pres chapter respectively.

Proof : Let $o < u < 1/q$. Then $u = \sum_{i=1}^{\infty} U_i q^{-i}$ and so u can be identi with the sequence $U = U_1U_2U_3\cdots$. If U is less than or equal to all its proper suffixes, we take $t = w = u$. Since $\bar{U}$ is greater t all its proper suffixes $s(U) = \log \rho$ where $\rho = \max_i \left\{ | \rho_i | \right\}$. $1-V(z) = \prod_i (1- \rho_i z)$ and $V(z) = \sum V_i z^i$ with $\bar{U} = V_1V_2V_3\cdots$. It is known that the complex reciprocal root of greatest magnitu is real and positive. Now $r(u) = -\log X = \log \frac{1}{X}$ where X is the unique positive real root of the equation $1-V(z) = 0$. Since $\frac{1}{X}$ the result follows. If U is not less than or equal to all its proper suffixes, let i be the least integer for which there exi

finite j such that $U_k = U_{k+i}$ for $1 \leq k \leq j$, but $U_{j+i} < U_j$.

Consider the polynomial

$$f(z) = \sum_{k=1}^{i-1} U_k z^k ,$$

and the analytic functions

$$f(z) = f(z) + (U_1-1) z^i + \frac{(q-1)z^{i+1}}{1-z} ,$$

and $$\phi(z) = \frac{f(z)+U_1 z^i}{1-z^i} .$$

Let $t = r(1/q) = .U_1U_2\ldots U_{i-1}(U_i-1)(\dot{q}-1)$ and

$w = \phi(1/q) = .\dot{U}_1U_2\ldots \dot{U}_i$.

Note that the expansions of both t and w are periodic, The expansion of t ends in an infinite sequence of $(q-1)$'s, while w is purely periodic with period i. Clearly $U_1 \neq 0$ and $t \leq u \leq w$. Since t and w are both less than or equal to all their proper suffixes and $s(u)$ is monotonically increasing function,

$$r(t) = s(t) \leq s(u) \leq s(w) = r(w).$$

By (4) and (5) $q^{-r(t)}$ is the real positive root of the equation

$$f(z) + (U_1-1)z^i + \frac{(q-1)z^{i+1}}{1-z} = \frac{qz-1}{1-q}$$

or $$(1-z)f(z) = qz-1-(q-U_1)z^{i+1}-(U_1-1)z^i \qquad \ldots(29)$$

and $q^{-r(w)}$ is the real positive root of the equation

$$\frac{f(z)-u_1 z^1}{1-z^1} = \frac{g(z)-1}{1-z}$$

or $(1-z)f(z) = g(z)-1-(g-u_1)z^1 + \ldots -(u_1-1)z^1$.

Since (2) has a unique positive real root, $q^{-r(t)} = q^{-r(w)}$. Hence $r(t) = r(w)$. This completes the proof.

Theorem 3 : If $s(u) = 1 - \frac{\epsilon}{\ln q^q}$ and ϵ is small then

$$u = \left(\frac{\epsilon}{q-1}\right)^{1 + \frac{\epsilon}{\ln q^q}} \left(1 + \frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)h(u-O(\epsilon))}{\ln q^q} + O\left(\frac{\epsilon}{q-1}\ln\frac{\epsilon}{q-1}\right)\right)\right)$$

and

$$u = \left(\frac{\epsilon}{q-1}\right)^{1 + \frac{\epsilon}{\ln q^q}} \left(1 + \frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)h(u-O(\epsilon))}{\ln q^q} + O\left(\frac{\epsilon}{q-1}\ln\frac{\epsilon}{q-1}\right)\right)\right)$$

Proof. As $s(u) \to 1$, $u \to 0$. So by theorem 2 there exist t and w such that $t \leq u \leq w$ and $s(u) = r(t) = r(w)$ and by theorem 1

$$t = \left(\frac{\epsilon}{q-1}\right)^{1 + \frac{\epsilon}{\ln q^q}} \left(1 + \frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)h(t)}{\ln q^q} + O\left(\frac{\epsilon}{q-1}\ln\frac{\epsilon}{q-1}\right)\right)\right)$$

$$w = \left(\frac{\epsilon}{q-1}\right)^{1 + \frac{\epsilon}{\ln q^q}} \left(1 + \frac{\epsilon}{q-1}\left(\frac{q+1}{2q} - \frac{(q-1)h(w)}{\ln q^q} + O\left(\frac{\epsilon}{q-1}\ln\frac{\epsilon}{q-1}\right)\right)\right)$$

Since $w-t = O(\epsilon)$, $t = u-O(\epsilon)$ and $w = u + O(\epsilon)$.

By the above theorem as $u \to 0$ be can thus estimate

$$1-s(u) = \frac{u}{\ln q^q} \qquad \ldots (30)$$

We know from chapter 3 that

$$I(q^m-1,\ uq^m) \sim n^{s(u)}$$

or $\quad \sim n^{H(u)-1}$

or $\quad \ln n^{-1} \sim (1-H(u)) \ln n.$

Using (1), we have

$$\ln n^{-1} \sim \frac{d \ln n}{n \ln q^q} \sim \frac{d}{n} \frac{\ln n}{\ln q^q}$$

or $$d \sim \frac{n \ln n^{-1}}{\log n}$$

Thus we have proved the following theorem.

Theorem 4. Long primitive, narrow-sense BCH codes over GF(q) have relative distance

$$\frac{d_d}{n} \sim \frac{n \ln n^{-1}}{\log n}$$

A family of codes over GF(q) is said to be <u>good</u> if it contains an infinite subsequence of codes $C_1, C_2, \ldots$, C_i an (n_i, k_i) code with distance d_i and both the ratios $R_i = k_i/n_i$ and d_i/n_i bounded away from zero as a limit as $i \to \infty$.

Primitive BCH codes do not have this property. For, consider an infinite sequence of primitive, narrow-sense BCH codes $C_1, C_2, \ldots$, C_i is an (n_i, k_i) code with distance d_i and fixed information rate $R = R_i = k_i/n_i$ for all $i = 1, 2, 3, \ldots$. Then by theorem 4 as $n \to \infty$,

$$\frac{d_d}{n} \sim \frac{q \ln n^{-1}}{\log n}$$

[illegible] is a fixed quantity and log n is an [illegible] of n, d√n -> ∞.

In view of the above result, we should search for good codes in other types of cyclic codes. There is a similar question : Are cyclic codes over GF(, good ?

REFERENCES

[1] E. R. Berlekamp, the enumeration of information symbols in BCH codes, Bell System Technical Journal 46,(1967b), 1861-1880.

[2] E.R. Berlekamp, long primitive binary BCH codes have distance $d \sim 2n \ln R^{-1}/\log_2 n$, IEEE Information Theory 18, (1972), 415-426.

[3] Ian F. Blake, Algebraic Coding Theory : History and Development, Dowden Hutchinson and Ross, 1973.

[4] R. C. Bose and D.K. Ray - Chaudhuri, on a class of error correcting binary group codes, Information and Control 3, (1960), 68-79.

[5] R.C. Bose and D.K. Ray-Chaudhuri, further results on error correcting binary group codes, Information and Control 3,(1960) , 279-290.

[6] M.J.E. Golay, notes on digital coding, Proc. IRE 37, (1949), 657.

[7] D.C. Gorenstein and N. Zierler, a class of error correcti codes in p^m symbols, Journal of Social and Industrial Applied Mathematics 9, (1961), 207-214.

[8] R. W. Hamming, error detecting and error correcting codes, Bell System Technical Journal. 29 (1950), 147-160.

[9] B. Mann Henry, On number of information symbols in Bose-Chaudhuri codes, Information and Control 5,(1963),153-16

[10] S. Lin and E. J. Weldon Jr., long BCH codes are bad, Information and Control 11, (1967), 445-451.

[11] W. W. Peterson, encoding and error correcting procedures for Bose and Chaudhuri codes, I.R.E. Trans, Information Theory 6, (1960), 459-470.

[12] W. W. Peterson and E.J. Weldon Jr., Error Correcting Codes, MIT Press, 1972.

[13] I.S. Reed and G. Solomon, polynomial codes over certai finite fields, Journal of Social and Industrial Applie Mathematics 8, (1960), 300-304.

[14] D. Slepian , a class of binary signalling alphabets, Bell System Technical Journal 35, (1956), 203-234.